# An end to end image processing pipeline which removes glare and lens flare

Team members: Noa Glaser

## Motivation

Broadly, using machine learning to recover information that lost in a scene is very interesting. Some examples of information that may be lost, corrupted, or unavailable in an image can be pixels saturated or distorted due to glare or lens flare, and occlusions due to scene geometry (ie when generating novel views based on the 3D geometry recovered from a light field).

In this project we will focus on correcting for lens flares and excessive glare. These effects, while sometimes used artistically, can be very frustrating and create unwanted artifacts. Some hardware stacks might be more prone to these effects. For example, when the Google Pixel first launched, it received a lot of criticism as the design of the lens stack led to excessive lens flares. Hardware is extremely expensive or impossible to change, but software is not. Therefore, this solution can also be useful to camera manufacturers encountering similar issues.

## Related Work

There are two high level approaches to recovering information from images.

The first is simply making this information up. More formally known as pixel in-painting, one can use either priors such as information from surrounding pixels, patches in the image, or a trained neural network to “hallucinate” the missing information. For example, (Eilertsen, Kronander, Denes, Mantiuk, & Unger, 2017) showed that they can recover the blown out parts of an HDR image using a convolutional neural net trained on previous images.

The second approach is to use actually recover the information through some interesting encoding of the unseen information in the scene. Through very controlled environments, or capturing a video sequence, (Saunders, Murray-Bruce, & Goyal, 2019) and (Bouman & etal, 2017) demonstrated that you can recover information about occluded objects from their penumbras (shadows they cast by other objects.) Therefore, there is some hope that objects that are no longer directly visible to the sensor will have information encoded in other parts of the image.

If there is information hidden in the scene, it might be easier to recover from the raw camera output, before it is processed by the ISP to be demosaiced, denoised etc. For example, (Chen, Chen, Xu,, & Koltun, 2018) showed that they can capture very dimly lit scenes extremely well when processing the raw Bayer readout. Happily, (Hasinoff, et al., 2018) released a large dataset of raw images to be converted to HDR. Unfortunately those images are all underexposed and so might not contain training information with blown out pixels. Other datasets of raw images include RAISE (Dang-Nguyen, Pasquini, Conotter, & Boato, 2015) and FiveK (Bychkovsky, Paris, Chan, & Durand, 2011).

Glare, reflection and lens flare removal solutions are of considerable interest and have been implemented in commercial applications. Google's HDR+ mode seems to subdue the lens flares which plagued its devices (Chavez, 2016) and its photo scanning application merges multiple views to remove glare from photo lamination (Liu, Rubinstein, Krainin, & Freeman, 2017).

Though this project will focus purely on software, it should be noted that hardware can be very useful for avoiding or recovering from glare, lens flare, or pixel saturation in general. Removing the effects of glare and reflections is usually approached with polarizing filters. (Rouf, Trentacoste, Mantiuk, C., & Heidrich, 2011) engineered the "bokeh" point spread function of a lens such that an oversaturated pixel spread out its values to neighboring pixels, thereby encoding its information in their values.

Approach

Our solution will be an end to end "programmable ISP" that takes in raw images and outputs processed images free of excessive glare and lens flares. The ISP will be a neural network potentially followed by some additional white balancing/global tone mapping component (as (Chen, Chen, Xu,, & Koltun, 2018) showed that their ISP model struggled with global tone mapping). We will experiment with a U-net architecture, similar to (Chen, Chen, Xu,, & Koltun, 2018), and incorporate a CNN architecture similar to (Eilertsen, Kronander, Denes, Mantiuk, & Unger, 2017) as the legs of the CNN.

In regards to dataset, we will need a lot of RAW images with unsavory lens flare and glare effects. Unfortunately, most of the raw image datasets (Hasinoff, et al., 2018), (Dang-Nguyen, Pasquini, Conotter, & Boato, 2015), (Bychkovsky, Paris, Chan, & Durand, 2011) consist of curated, nice images. A high quality dataset for validating the algorithm might need to be gathered by taking a lot of images with a hooded and un-hooded DSLR. In order to get started, we can synthesize a database by using one of the many applications which add lens flares to existing images and converting the modified images back to a Bayer array.

Milestones

Week 1 - Feb 17-23: Assemble datasets

Week 2 - Feb 24-Mar 2: Validate simple model (overtrain CNN), continue assembling data

Week 3 – Mar 3 – Mar 9: Iterate on models, train

Week 4 – Mar 10-Mar 15: Time profile models, assemble poster/paper, clean up codebase

## Works Cited

Bouman, K. L., & etal. (2017). Turning Corners into Cameras: Principles and Methods. *ICCV*. Retrieved from https://people.csail.mit.edu/klbouman/pw/papers_and_presentations/cornercam_iccv2017.pdf

Bychkovsky, V., Paris, S., Chan, E., & Durand, F. (2011). Learning Photographic Global Tonal Adjustment with a Database of Input / Output Image Pairs. The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition.

Chavez, C. (2016, December 9). *Google Pixel's New Camera Update Attempts to Fix Lens Flare*. Retrieved from phandroid.com: https://phandroid.com/2016/12/09/google-pixel-camera-lens-flare-halo-effect-fix-update/

Chen, C., Chen, Q., X. J., & Koltun, V. (2018). Learning to See in the Dark. *CVPR*. Retrieved from https://arxiv.org/abs/1805.01934

Dang-Nguyen, D.-T., Pasquini, C., Conotter, V., & Boato, G. (2015, March 18). RAISE – A Raw Images Dataset for Digital Image Forensics. Portland, Oregon: ACM Multimedia Systems. Retrieved from http://loki.disi.unitn.it/RAISE/

Eilertsen, G., Kronander, J., Denes, G., Mantiuk, R. K., & Unger, J. (2017). HDR image reconstruction from a single exposure using deep CNNs. *36(6)*. ACM Transactions on Graphics (TOG). Retrieved from http://hdrv.org/hdrcnn/

Hasinoff, S. W., Sharlet, D., Geiss, R., Adams, A., Barron, J. T., Kainz, F., . . . Levoy, M. (2018, February 15). HDR+ Burst Photography Dataset - Details. *HDR+ Burst Photography Dataset - Details*. Mountain View, CA, USA: Google Research. Retrieved from https://hdrplusdata.org/dataset.html

Liu, C., Rubinstein, M., Krainin, M., & Freeman, B. (2017, April 20). *PhotoScan: Taking Glare Free Pictures of Pictures*. Retrieved from Google AI Blog: https://ai.googleblog.com/2017/04/photoscan-taking-glare-free-pictures-of.html

Rouf, M., Trentacoste, M., Mantiuk, R., C., L., & Heidrich, W. (2011). Glare encoding of high dynamic range images. *CVPR*, 289-296. doi:10.1109/CVPR.2011.5995335

Saunders, C., Murray-Bruce, J., & Goyal, V. K. (2019, January). Computational Periscopy With An Ordinary Digital Camera. *Nature*(7740), 472-475. doi:10.1038/s41586-018-0868-6